

44

**In the United States Patent and Trademark Office**

Application Number: 09/934,156  
Filing Date: 08/21/2001  
First Named Inventor: David Roth Rigney  
Application Title: System, Methods, and Computer Program Product  
for Analyzing Microarray Data  
Group Art Unit: 2176

Mailed: 11/19/2001  
Austin TX

**Information Disclosure Statement**

Assistant Commissioner for Patents  
Washington, D.C. 20231

Dear Commissioner:

Please find attached a completed Form PTO-1449 (substitute PTO/SB/08 A&B), along with copies of the pertinent parts of the references cited on the Form. All of the references are in English.

The following are comments on the relevance of the references. The page numbers mentioned below are those of the patent application. The cite numbers correspond to the order of appearance of the references in the application.

(Cite No. 1) DUGGAN et al. is cited on page 2. It is a review of the technical aspects of cDNA microarrays, including general principles, fabrication of the arrays, target labelling, image analysis and data extraction, management and mining.

(Cite No. 2) CHEN et al. is cited on page 3. This patent describes methods for estimating the ratio of data corresponding to each microarray spot (intervention value/control value). All such preliminary calculations, which compensate for errors introduced by the measurement process itself, are assumed to have been applied before using my methods.

(Cite No. 3) NEWTON et al. is cited on pages 3 and 28. This manuscript describes methods for estimating the ratio of data corresponding to each microarray spot (intervention value/control value). All such preliminary calculations, which compensate for errors introduced by the measurement process itself, are assumed to have been applied before using my methods.

(Cite No. 4) IYER et al. is cited on pages 3, 5, 8, 14, 30, 49, 51, 57, 58, 59, 61, 62 and 63. This publication illustrates the collection of microarray data and the application of one clustering algorithm to those data, namely, the algorithm described in EISEN et al. The data described in this publication are publicly available, and I used them to demonstrate the methods of the patent application.



(Cite No. 5) EISEN et al. is cited on pages 4, 5, 49 and 62. This publication describes the application of a general-purpose clustering algorithm to time-series microarray data.

(Cite No. 6) TAVAZOIE et al. is cited on pages 4, 5, 7 and 8. This publication describes the application of a general-purpose clustering algorithm to time-series microarray data. The authors also use a database containing a functional classification of yeast genes, in order to determine whether the clustering results in groupings (clusters) that are preferentially associated with particular functions.

(Cite No. 7) TAMAYO et al. is cited on pages 4 and 5. This publication describes the application of a general-purpose clustering algorithm to time-series microarray data.

(Cite No. 8) BEN-DOR et al. is cited on pages 4 and 5. This publication describes the application of a general-purpose clustering algorithm to time-series microarray data.

(Cite No. 9) GETZ et al. is cited on pages 4 and 5. This publication describes the application of a general-purpose clustering algorithm to time-series microarray data.

(Cite No. 10) ZHENG et al. is cited on pages 4 and 5. This publication describes the application of general-purpose clustering algorithms to time-series microarray data.

(Cite No. 11) PEROU et al. is cited on page 4. This publication describes the application of a general-purpose clustering algorithm to microarray data corresponding to different types of interventions and types of cells.

(Cite No. 12) TIBSHIRANI et al. is cited on pages 4 and 5. This publication describes the application of a general-purpose clustering algorithm to microarray data corresponding to different types of cells.

(Cite No. 13) KAUFMAN and Rousseeuw is cited on pages 5 and 29. This book reviews general-purpose clustering algorithms and implements some of them with actual listings of computer programs. Chapter 3 describes a computer program for the clustering of large data sets (CLARA), which is incorporated into software described in my patent application.

(Cite No. 14) HOLMES et al. is cited on page 6. This conference proceeding describes an algorithm for finding clusters of genes that simultaneously have similar expression patterns in microarray data, as well as similar promoter DNA sequences. The rationale of the algorithm is that a particular transcription factor should recognize and bind to the promoters that are shared by genes in a cluster, which will consequently result in similar expression patterns for the genes in that cluster. This is the same rationale that I take, but instead of looking for the promoter sequence motif corresponding to a cluster, I look for the actual transcription rate that is modulated by the cluster's transcription factor, which is estimated from the microarray data (see pages 29-36 of the patent application).

(Cite No. 15) YEUNG et al. is cited on page 7. This manuscript describes methods for selecting among the many clustering algorithms that may be applied to microarray data. Whereas this manuscript selects among the algorithms using analysis of the microarray data themselves, the method in my patent application is to seek support for an algorithm based on a quantitative analysis of the content of literature about genes in each cluster (the cluster being generated through use of any algorithm), in particular, through the calculation of literature-based figure-of-merit indices.

(Cite No. 16) SHATKAY et al. is cited on page 10. This conference proceeding describes a method for finding words that characterize genes represented in a microarray. In contrast to my methods, it does not make use of the actual clustering of microarray data, and therefore it does not attempt to find the words and phrases that best distinguish each cluster from all the other clusters. It does not provide a method for selecting among clustering algorithms. It is also semi-automatic, rather than automatic, in that it requires an investigator to manually select some single publication that characterizes a gene (a "kernel" document).

(Cite No. 17) MASYS et al. is cited on page 10. This publication describes a method for finding words that characterize genes that have been clustered from microarray data. It does not attempt to find the words and phrases that best distinguish each cluster from all the other clusters. Consequently, as noted on page 326 of the publication, it can generate uninformative, commonly occurring words and phrases such as "amino acid sequence". Also, it does not provide a method for selecting among clustering algorithms.

(Cite No. 18) DELORIE is cited on page 16. This is a web page that describes the DJGPP development system that is used to put a unix-like software system on a Windows workstation.

(Cite No. 19) McCALLUM is cited on pages 18, 20, 37 and 39. This is a web page that describes the Rainbow computer program. Pages 10-20 list the command lines that Rainbow can execute.

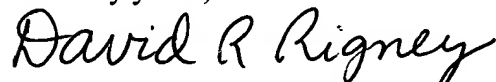
(Cite No. 20) PRESS et al. is cited on pages 24, 29, 32, 42 and 44. This book is a standard reference about numerical analysis algorithms. The system, methods, and computer program described in the patent application make use of algorithms described in Chapter 4 (Integration of Functions), Chapter 7 (Random Numbers), Chapter 8 (Sorting), Chapter 9 (Root Finding), and Chapter 14 (Statistical Description of Data).

(Cite No. 21) MANNING and SCHUETZE is cited on page 33. Chapter 14 of this book has a discussion of the EM algorithm in connection with the subject of clustering.

(Cite No. 22) MITCHELL is cited on page 42. Chapter 6 of this book has a discussion of Naive Bayes classification.

(Cite No. 23) EGGHE & ROUSSEAU is cited on page 46. Section 4 in chapter III of this book has a discussion of bibliographic coupling and co-citation analysis.

Sincerely yours,



David R. Rigney, Applicant Pro Se  
GENETWORKS Inc.

P.O. Box 33296

Austin TX 78764-0296 U.S.A.

Tel. (512) 445-7301

FAX (512) 445-7301

Email: drigney@genetworks.com

+

+

U.S. Patent and Trademark Office: U.S. DEPARTMENT OF COMMERCE

U.S. Patent and Trademark Office: U.S. DEPARTMENT OF COMMERCE

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.



Substitute for form 1449A/PTO

## INFORMATION DISCLOSURE STATEMENT BY APPLICANT

(use as many sheets as necessary)

Sheet

1

of

1

**Complete if Known**

Application Number	09/934,156
--------------------	------------

<b>Filing Date</b>	08/21/2001
--------------------	------------

First Named Inventor	David Rigney
----------------------	--------------

Group Art Unit	2176
----------------	------

Examiner Name \_\_\_\_\_

Attorney Docket Number

## U.S. PATENT DOCUMENTS

[illegible]

## FOREIGN PATENT DOCUMENTS

[illegible]

Examiner  
Signature

Date \_\_\_\_\_

**Considered**

\*EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to applicant.

<sup>1</sup> Unique citation designation number. <sup>2</sup> See attached Kinds of U.S. Patent Documents. <sup>3</sup> Enter Office that issued the document, by the two-letter code (WIPO Standard ST.3). <sup>4</sup> For Japanese patent documents, the indication of the year of the reign of the Emperor must precede the serial number of the patent document. <sup>5</sup> Kind of document by the appropriate symbols as indicated on the document under WIPO Standard ST. 16 if possible. <sup>6</sup> Applicant is to place a check mark here if English language Translation is attached.

**Burden Hour Statement:** This form is estimated to take 2.0 hours to complete. Time will vary depending upon the needs of the individual case. Any comments on the amount of time you are required to complete this form should be sent to the Chief Information Officer, U. S. Patent and Trademark Office, Washington, DC 20231. **DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. SEND TO:** Assistant Commissioner for Patents, Washington, DC 20231.



Please type a plus sign (+) inside this box → ☐

PTO/SB/08B (08-00)

Approved for use through 10/31/2002. OMB 0651-0031

U. S. Patent and Trademark Office: U.S. DEPARTMENT OF COMMERCE

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

Substitute for form 1449B/PTO

## INFORMATION DISCLOSURE STATEMENT BY APPLICANT

(use as many sheets as necessary)

Sheet 1 of 2

### Complete if Known

Application Number	09/934,156
Filing Date	08/21/2001
First Named Inventor	David Rigney
Group Art Unit	2176
Examiner Name	
Attorney Docket Number	

### OTHER PRIOR ART – NON PATENT LITERATURE DOCUMENTS

Examiner Initials*	Cite No. <sup>1</sup>	Include name of the author (in CAPITAL LETTERS), title of the article (when appropriate), title of the item (book, magazine, journal, serial, symposium, catalog, etc.), date, page(s), volume-issue number(s), publisher, city and/or country where published.	T <sup>2</sup>
	1	DUGGAN, D.J., M. Bittner, Y. Chen, P. Meltzer and J.M. Trent. Expression profiling using cDNA arrays. Nature Genetics 21, Suppl 1: 10-14 (1999).	
	3	NEWTON, M.A., C.M. Kendzierski, C.S. Richmond, F.R. Blattner, and K.W. Tsui. On differential variability of expression ratios. Tech. Rept. 139, Dept. of Biostatistics, Univ. of Wisconsin-Madison. <a href="http://www.stat.wisc.edu/~newton/papers/publications">http://www.stat.wisc.edu/~newton/papers/publications</a> (1999).	
	4	IYER, V.R., M.B. Eisen, D.T. Ross, G. Schuler, T. Moore, J.C.F. Lee, J.M. Trent, L.M. Staudt, J. Hudson Jr., M.S. Boguski, D. Lashkari, D. Shalon, D. Botstein and P.O. Brown. The transcriptional program in the response of human fibroblasts to serum. Science 283: 83-87 (1999).	
	5	EISEN, M.B, P.T. Spellman, P.O. Brown and D. Botstein. Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci USA 95: 14863-14868 (1998).	
	6	TAVAZOIE, S., J.D. Hughes, M.J. Campbell, R.J. Cho, and J.M. Church. Systematic determination of genetic network architecture. Nature Genetics 22:281-285 (1999).	
	7	TAMAYO, P, D. Slonim, J. Mesirov, Q. Zhu, S. Kitareewan, E. Dmitrovsky, E.S. Lander, and T.R. Golub. Interpreting patterns of gene expression with self-organizing maps: methods and application to hematopoietic differentiation. Proc Natl Acad Sci USA 96: 2907-2912 (1999).	
	8	BEN-DOR, A., R. Shamir and Z. Yakhini. Clustering gene expression patterns. J. Computational Biology 6:281-297 (1999).	
	9	GETZ, G., E. Levine, E. Domany, and M.Q. Zhang. Super-paramagnetic clustering of yeast gene expression profiles. arXiv:physics/9911038 at preprint server xxx.lanl.gov (1999).	
	11	PEROU, C.M., S.S. Jeffrey, M. van de Rijn, C.A. Rees, M.B. Eisen, D.T. Ross, A. Pergamenschikov, C.F. Williams, S.X. Zhu, D. et al. Distinctive gene expression patterns in human mammary epithelial cells and breast cancers. Proc Natl Acad Sci USA 96: 9212-9217 (1999).	
	12	TIBSHIRANI, R., T. Hastie, M. Eisen, D. Ross, D. Botstein, and P. Brown. Clustering methods for the analysis of DNA microarray data. Manuscript from Web site <a href="http://www-stat.stanford.edu/~tibs/lab/publications.html">http://www-stat.stanford.edu/~tibs/lab/publications.html</a> (1999).	
	13	KAUFMAN, L. and P.J. Rousseeuw. Finding Groups in Data: An Introduction to Cluster Analysis. New York: Wiley (1990)	

Examiner  
Signature

Date  
Considered

\*EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to applicant.

<sup>1</sup> Unique citation designation number. <sup>2</sup> Applicant is to place a check mark here if English language Translation is attached.

Burden Hour Statement: This form is estimated to take 2.0 hours to complete. Time will vary depending upon the needs of the individual case. Any comments on the amount of time you are required to complete this form should be sent to the Chief Information Officer, U. S. Patent and Trademark Office, Washington, DC 20231. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. SEND TO: Assistant Commissioner for Patents, Washington, DC 20231.

Please type a plus sign (+) inside this box → ☐

PTO/SB/08B (08-00)

Approved for use through 10/31/2002. OMB 0651-0031

U. S. Patent and Trademark Office: U.S. DEPARTMENT OF COMMERCE

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

Substitute for form 1449B/PTO

## INFORMATION DISCLOSURE STATEMENT BY APPLICANT

(use as many sheets as necessary)

Sheet 2 of 2

### Complete if Known

Application Number	09/934,156
Filing Date	08/21/2001
First Named Inventor	David Rigney
Group Art Unit	2176
Examiner Name	
Attorney Docket Number	

### OTHER PRIOR ART - NON PATENT LITERATURE DOCUMENTS

Examiner Initials*	Cite No. <sup>1</sup>	Include name of the author (in CAPITAL LETTERS), title of the article (when appropriate), title of the item (book, magazine, journal, serial, symposium, catalog, etc.), date, page(s), volume-issue number(s), publisher, city and/or country where published.	T <sup>2</sup>
	14	HOLMES, I. and W.J. BRUNO. Finding regulatory elements using joint likelihoods for sequence and expression profile data. Proc. Eighth International Conference on Intelligent Systems for Molecular Biology, AAAI Press, Menlo Park CA, pp. 202-210. (2000).	
	15	YEUNG, K.Y., D.R. Haynor and W.L. Ruzzo. Validating clustering for gene expression data. Technical report UW-CSE-00-01-01. University of Washington, Seattle. Manuscript from Web site <a href="http://www.cs.washington.edu/homes/kayee/research.html">http://www.cs.washington.edu/homes/kayee/research.html</a> (2000).	
	16	SHATKAY, H., S. Edwards, W.J. Wilbur, and M. Boguski. Genes, themes, and microarrays. Using information retrieval for large-scale gene analysis. Proc. Eighth International Conference on Intelligent Systems for Molecular Biology, AAAI Press, Menlo Park CA, pp. 317-328. (2000).	
	17	MASYS, D.R., J.B. Welsh, J.L. Fink, M. Gribskov, I. Kiacansky, and J. Corbeil. Use of keyword hierarchies to interpret gene expression patterns. Bioinformatics 17: 319-326 (2001).	
	18	DELORIE, D.J. Guide: What is DJGPP? Web page from <a href="http://www.delorie.com/djgpp/doc/ug/intro/what-is-djgpp.html">http://www.delorie.com/djgpp/doc/ug/intro/what-is-djgpp.html</a> (1997).	
	19	McCALLUM, A.K. Rainbow. Web page from <a href="http://www.cs.cmu.edu/~mccallum/bow">http://www.cs.cmu.edu/~mccallum/bow</a> (1998).	
	20	PRESS, W.H., S.A. Teukolsky, W.T. Vetterling, and B.P. Flannery. Numerical Recipes in C. The Art of Scientific Computing, Second. edn. New York: Cambridge University Press (1992).	
	21	MANNING, C.D. and H. Schutze. Foundations of Statistical Natural Language Processing. Cambridge MA: The MIT Press (1999).	
	22	MITCHELL, T. Machine Learning. New York: McGraw-Hill (1997).	
	23	EGGHE, L. and R. Rousseau. Introduction to Informetrics. Quantitative Methods in Library, Documentation, and Information Science. Amsterdam: Elsevier Science Publishers (1990).	

Examiner  
Signature

Date  
Considered

\*EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to applicant.

<sup>1</sup> Unique citation designation number. <sup>2</sup> Applicant is to place a check mark here if English language Translation is attached.

Burden Hour Statement: This form is estimated to take 2.0 hours to complete. Time will vary depending upon the needs of the individual case. Any comments on the amount of time you are required to complete this form should be sent to the Chief Information Officer, U. S. Patent and Trademark Office, Washington, DC 20231. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. SEND TO: Assistant Commissioner for Patents, Washington, DC 20231.